

A tiered approach to estimate inventory data and impacts of chemical products and mixtures

Gregor Wernet · Stefanie Hellweg ·
Konrad Hungerbühler

Received: 28 February 2011 / Accepted: 27 February 2012 / Published online: 13 March 2012
© Springer-Verlag 2012

Abstract

Purpose Mixtures of organic chemicals are a part of virtually all life cycles, but LCI data exist for only relatively few chemicals. Thus, estimation methods are required. However, these are often either very time-consuming or deliver results of low quality. This article compares existing and new methods in two scenarios and recommends a tiered approach of different methods for an efficient estimation of the production impacts of chemical mixtures.

Methods Four approaches to estimate impacts of a large number of chemicals are compared in this article: extrapolation from existing data, substitution with generic datasets on chemicals, molecular structure-based models (MSMs, in this case the Finechem tool), and using process-based estimation methods. Two scenarios were analyzed as case studies: soft PVC plastic and a tobacco flavor, a mixture of 20 chemicals.

Responsible editor: Guido Sonnemann

Electronic supplementary material The online version of this article (doi:10.1007/s11367-012-0404-0) contains supplementary material, which is available to authorized users.

G. Wernet (✉)
The ecoinvent Centre,
Empa, Ueberlandstrasse 129,
8600 Dübendorf, Switzerland
e-mail: wernet@ecoinvent.org

S. Hellweg
Institute for Environmental Engineering, ETH Zurich,
Wolfgang-Pauli-Strasse 15,
8093 Zurich, Switzerland

K. Hungerbühler
Institute for Chemical and Bioengineering, ETH Zurich,
Hönggerberg, HCI,
8093 Zurich, Switzerland

Results Process models have the potential to deliver the best estimations, as existing information on production processes can be integrated. However, their estimation quality suffers when such data are not available and they are time-consuming to apply, which is problematic when estimating large numbers of chemicals. Extrapolation from known to unknown components and use of generic datasets are generally not recommended. In both case studies, these two approaches significantly underestimated the impacts of the chemicals compared to the process models. MSMs were generally able to estimate impacts on the same level as the more complex process models. A tiered approach using MSMs to determine the relevance of individual components in mixtures and applying process models to the most relevant components offered a simpler and faster estimation process while delivering results on the level of most process models.

Conclusions The application of the tiered combination of MSMs and process models allows LCA practitioners a relatively fast and simple estimation of the LCIA results of chemicals, even for mixtures with a large number of components. Such mixtures previously presented a problem, as the application of process models for all components was very time-consuming, while the existing, simple approaches were shown to be inadequate in this study. We recommend the tiered approach as a significant improvement over previous approaches for estimating LCA results of chemical mixtures.

Keywords Chemical production · Finechem · Flavor · Inventory estimation · LCA

1 Introduction

Organic chemicals are a particular concern in life cycle assessment due to their presence in the life cycles of virtually all

industrial and consumer products. Consumer goods mostly composed of organic chemical products, such as detergents, pigments, or plastics, are important product classes, but even beyond the obvious importance of chemicals in these products, organic chemicals are vital during the manufacture of most other goods. Life cycle inventories of organic chemicals are therefore a part of every LCA and a feasible and accurate approach to setting up chemical inventories are important for the LCA community as a whole.

A range of chemical production inventories is available from database initiatives (e.g., NREL 2008; The ecoinvent Centre 2008a) or industry associations, e.g., the frequently used PlasticsEurope datasets on basic chemicals (PlasticsEurope 2005). Nevertheless, the available inventories cover only a few hundred of the many thousands of chemicals in production today. Most of these are very basic chemicals, while only few detailed studies on fine and advanced chemicals exist (Jiménez-Gonzalez et al. 2004; Wernet et al. 2010). In addition to the sheer number of chemicals that still need to be assessed, the difficulty of acquiring production data on chemicals is also problematic. Chemical companies are highly sensitive about access to their production data, as these data may give competitors or customers unwanted insights into a company's core business. Forming average datasets over several producers can be a solution but is less practical for specialty products with few producers. While efforts of primary data collection are ongoing, it is therefore clear that for the vast majority of chemicals, estimation methods for inventory data are needed.

The established method of estimating LCA results of chemical productions is to model the gate-to-gate inventories of the individual process steps to the best of the knowledge available and to then combine those process step data into a full cradle-to-gate inventory, possibly making use of established databases to cover upstream processes (Geisler et al. 2004; Hirschier et al. 2004; Jiménez-Gonzalez et al. 2000). If some information on individual process steps is available, the remaining gaps are sometimes filled with models that cover specific parts of a process step. Models exist for parts of chemical reactions, including steam and energy use (Szijarto et al. 2008; Van der Vorst et al. 2009), solvent recovery (Capello et al. 2007), and waste water treatment (Köhler et al. 2007). Many of these models require detailed information about the production process. In addition, several chemical companies have in-house tools with which they model parts of their production (Bretz and Frankhauser 1997; Curzons et al. 2007; Saling et al. 2002). Obviously, the more actual on-site data on the process are available, the better. Also, some fallback solutions and standard estimations have been proposed for individual parts of chemical processes, such as those of Hirschier et al. (2004) and Jiménez-Gonzalez et al. (2000). On the other hand, if only the general synthesis route is known without process details, the whole LCI has to be estimated, either by completely relying on models and standard estimations or by using generic models on a process

step level (e.g., Geisler et al. 2004). This may be a valid option if general information on production is available, e.g., through documentation on industrial chemistry (Thieme Chemistry 1995; Ullmann's Encyclopedia of Industrial Chemistry 2006).

All the methods mentioned above follow the fundamental approach of creating individual process estimations and combining these individual units into a full cradle-to-gate inventory. This approach has two potential weaknesses. First of all, if many process steps are estimated, especially if standard estimations for basic chemicals are used while highly complex reactions are involved, the uncertainties will increase over the processes and may lead to prohibitively high overall uncertainties in the resulting inventories. As most models and standard estimations have been developed with data on basic chemical production, applying these estimation procedures and models to fine chemical productions, which are often more complex and less resource and energy efficient, may introduce systematic errors. And secondly, if the synthesis route is not known, these approaches are not possible. In order to palliate these shortcomings, molecular structure-based models (MSMs) were recently developed as an alternative approach in situations where the process-based approach fails (Wernet et al. 2008, 2009). These models can estimate cradle-to-gate LCA results in terms of cumulative energy demand (CED, VDI 1997), the global warming potential (GWP, IPCC 2007), and Ecoindicator 99 scores (EI99, Goedkoop and Spriensma 2000) based solely on the molecular structure. They therefore allow screening LCAs even if no data on the production of a chemical are available. However, also this approach is not without drawbacks. For instance, it is a little transparent and can be characterized as a "black box" approach. Further, this is a rather new approach and has so far not been widely applied in practice. A comparison of the different methods and approaches may hence be helpful to understand the merits and shortcomings of each of these approaches better and to provide a guideline on which method to use in a given situation. The objective of this article is to compare different data estimation techniques for chemicals and to identify an appropriate estimation procedure. Both the quality of the results and the time and effort required will be evaluated on two case studies of chemical products which exemplify the many different scenarios imaginable.

2 Materials and methods

2.1 Modeling approaches

In order to illustrate the application of the various methods to set up inventories of chemical productions, we compare them using two very different examples: soft polyvinylchloride (PVC) production and the production of a tobacco flavor. Both are examples of common scenarios where detailed inventory information on chemicals may be lacking. An LCA

practitioner assessing a consumer good containing such chemical mixtures therefore encounters a lack of data and needs to estimate the LCI or the production impacts. Soft PVC for example contains, on average, 30 % by weight plasticizers, such as DEHP (bis(2-ethylhexyl)phthalate, EU JRC 2008). Yet the ecoinvent database (The ecoinvent Centre 2008a), one of the most commonly used sources of LCI data worldwide, contains data on PVC production but not on the synthesis of the plasticizer. A practitioner therefore faces the question of how to assess the soft PVC with data on only 70 % of the product.

We identified several common solutions for abridging data gaps on chemical production and applied these to the case study. Neglecting the missing inventories completely is not a viable option, as this would guarantee that the final LCA underestimates the real impacts and therefore would lead to questionable results.

2.1.1 Extrapolation

The practitioner may neglect the chemical with no production data when determining the composition. In the PVC example, for instance, the PVC dataset of ecoinvent could be used to describe the soft PVC. In practice, this means using 1 kg of PVC as an approximation for the 0.7-kg PVC and 0.3-kg DEHP which are actually present in the soft PVC. This way, the missing inventories are not completely neglected but are substituted by the inventory which exists.

2.1.2 Approximation with similar chemicals

The practitioner may use another dataset to approximate the missing inventory data. This is a common solution for non-chemical products, but since small changes in the molecular structure may have drastic impacts on the synthesis route required, it can be very difficult to apply to chemicals. Often, it is not possible to identify a product with an existing LCI where productions are similar enough to be a good substitute. The ecoinvent database provides a dataset called “chemicals organic, at plant” to alleviate this problem. This dataset represents a production mix of 20 common basic chemicals. The practitioner may use this dataset to describe, for instance, the production of DEHP by creating a soft PVC inventory comprised of 0.7 kg PVC and 0.3 kg “chemicals organic.”

2.1.3 Molecular structure-based models

The practitioner may use an MSM, e.g., the Finechem tool (available at <http://www.sust-chem.ethz.ch/tools/finechem>; Wernet et al. 2009), to estimate the impacts of DEHP production. This option is quite novel, as Finechem was released only in the autumn 2009.

The Finechem tool is based on artificial neural networks (ANN), a computational model structure that is based on the

principles of neuron interaction in the brain (Swingler 1996). ANN models are useful for nonlinear regression and pattern recognition. They are adaptive systems, meaning that an ANN model needs to be *trained* with training data to *learn* about the underlying correlations and patterns between a given input and output. In the case of the Finechem models, the input is a set of molecular parameters based on the molecular structure called descriptors, e.g., the presence and number of various functional groups and elements, the molecular weight, etc. In this approach, the molecular structure is thus used as a basis to gain some insight on the conditions of the necessary production processes. The Finechem models were trained on a set of inventory data compiled from industry sources and the ecoinvent database. It is important to note that during the creation of the Finechem models, only inventories based on actual process data for individual chemical productions were used. The ecoinvent database also contains datasets based on the process models described below under IV, but these data were not used during the creation of the Finechem models. They are therefore no hidden links or correlations between the results of approaches III and IV. To avoid a *training bias*, a technique called cross-validation was used (Swingler 1996), so that the model quality was actually determined on the ability to accurately predict new chemicals, not chemicals that were included in the training set. Therefore, the available data were split into *training sets* and *test sets* (Wernet et al. 2009).

ANNs trained in such a way have acquired the ability to make predictions about chemicals that were not included in the training data. The application of the Finechem tool then works along the following steps: First, the user determines the descriptors for the chemical to be assessed. The Finechem tool uses ten descriptors. Then, the models calculate the output (e.g., the CED) based on the input of the descriptors. The underlying calculations of a neural network model can be followed in real time, but since the interactions between the individual neurons are very complex and usually too complex for a rational analysis, they are usually considered to be black box models. Each Finechem model is actually a set of 30 individual models that were trained in slightly different ways, and the result is based on the distribution of the 30 individual predictions.

Molecular structure-based models based on ANNs can be a very powerful tool to determine complex correlations between features of the molecular structure and production requirements of chemicals. The Finechem models represent the first implementation of this modeling approach in LCA. Their current limitations are mostly due to the lack of available training data for more capable models (Wernet et al. 2009). Due to their black box nature, the estimation uncertainty needs to be handled differently than in the case of process models (see Section 4 below).

2.1.4 Process models

The practitioner may choose to invest the time to generate an LCI using process models. If details of production techniques are available, detailed process models on the level of individual process steps and equipment usage such as (Szijarto et al. 2008) may be used. In most cases without the cooperation of the producer, however, a simple approximation based on standard estimations is used. These are based on typical chemical productions. When applying such standard estimations, consistency is important. In this work, we applied the parameters suggested in Hischier et al. (2004). These suggested, e.g., a yield of 95 % for an average chemical reaction, the use of 2 MJ of energy in the form of steam, and 0.333 kWh of electricity. Air emissions of volatile chemicals were estimated to be 0.2 % of the input, and water emissions were estimated with a 90 % reduction of organic material in waste water treatment based on Köhler et al. (2007).

2.1.5 Tiered approach

In cases where a chemical mixture is made up of many chemicals, the process models will be very time-consuming to apply. For such a case, we propose a tiered approach in the form of a combination of MSMs and process-based models. For this tiered approach, the Finechem tool is applied first as a screening tool to identify

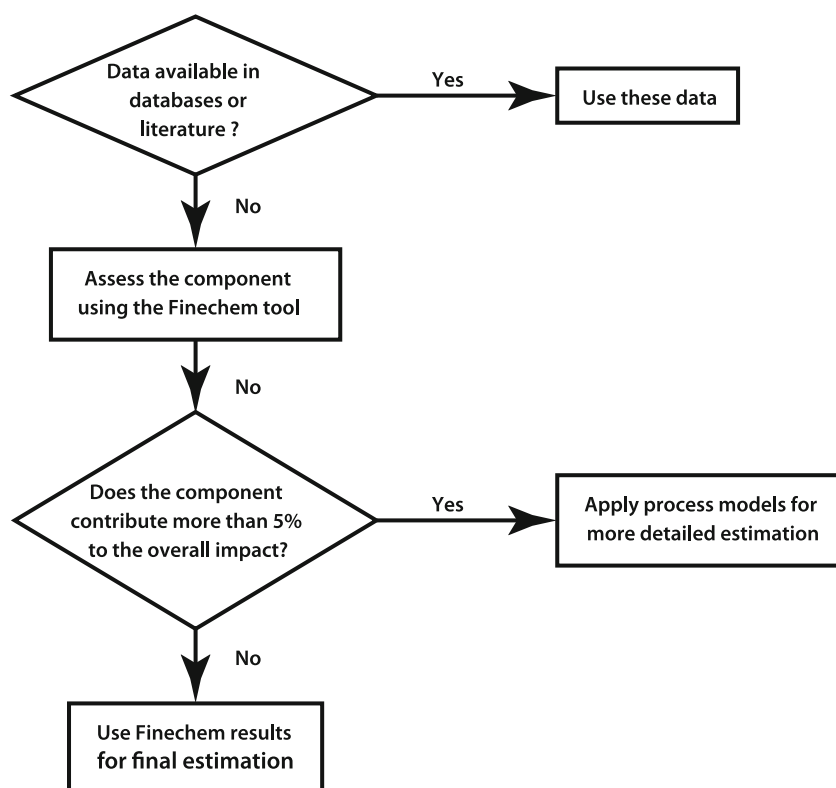
relevant chemicals for which further data refinement is suggested. If the impact of a chemical times its fraction in the mixture is larger than 5 % of the total impact according to the Finechem results, a process-based approach is applied to provide more detailed information on the most relevant chemicals. This approach is described in the flowchart in Fig. 1, and its application is demonstrated in the flavor case study.

The five approaches outlined above were used in both case studies, in order to illustrate the application and to compare the various methods.

2.2 Case study A: soft PVC

The PCV content of soft PVC (70 %) was always assessed using the available dataset from the ecoinvent database. Therefore, all results only differ due to the modeling approaches used to estimate the production impacts of the plasticizer DEHP. To model the production of DEHP on a process level (approach 4 above), knowledge of the synthesis route was necessary. This information was obtained from reference handbooks and chemical encyclopedias (Thieme Chemistry 1995; Ullmann's Encyclopedia of Industrial Chemistry 2006). According to these sources, DEHP is synthesized by esterification of phthalic anhydride with 2-ethylhexanol. Phthalic anhydride is produced through the catalytic oxidation of *ortho*-xylene, while 2-ethylhexanol is produced by aldol

Fig. 1 Flowchart for the tiered approach to estimating impacts of individual components of chemical mixtures



condensation of two molecules of butyraldehyde, followed by a hydrogenation. Butyraldehyde is a product of the hydroformylation of propylene. Some of these chemicals were covered by the ecoinvent database, but three reaction steps had to be estimated. The aldol reaction, the hydrogenation, and the following esterification had to be modeled using estimates. The synthesis of DEHP is described in Fig. 2.

The propylene hydroformylation process results in butanol after catalytic reduction of butyraldehyde. The dataset was adapted by removing the final hydrogenation step to create an LCI for butyraldehyde. Reaction stoichiometry and some yield data were taken from Ullmann's Encyclopedia of Industrial Chemistry (2006). No emission data and no data on energy use during the production were available. This is often the case, as even producers generally only gather emission and energy data on a plant scale, if at all (Szijarto et al. 2008). Therefore, in approach 4, the method described in Hischer et al. (2004) was used to fill the remaining data gaps on transport, emissions, and water and energy use. This is consistent with the approach chosen in several data gathering projects in ecoinvent (The ecoinvent Centre 2008b).

2.3 Case study B: tobacco flavor

In the second scenario, the production inventory of an artificial flavor was assessed. The composition of a tobacco flavor produced by Givaudan was extracted from a patent application (Naegeli 1978). The flavor consisted of 20 components, all of them organic chemicals ranging from basic to advanced. Four of these substances were covered by the ecoinvent database, and propyl acetate could be assessed with the dataset of the very similar isopropyl acetate. These five chemicals were assessed based on the process data available in ecoinvent. Since actual production data were

available, it was therefore unnecessary to model the impacts of these chemicals with the approaches above. For the remaining 15 chemicals, no data were available in the ecoinvent database. All 20 chemicals were within the application range of the Finechem tool (Wernet et al. 2009). According to the patent, 2–10 g of this mixture is applied to 500 g of fresh cut tobacco. Table 1 shows the composition of the flavor.

Several of the 15 chemicals for which no ecoinvent inventories are available were quite complex. This case study thus represents the rather common situation where the LCA practitioners are faced with a potentially time-consuming modeling task. For these 15 chemicals, the modeling approaches 1–5 were applied.

3 Results

3.1 Soft PVC

Application of approaches 1–3 was simple and straightforward (Fig. 3). The data basis for the generation of the process inventories according to approach 4 above (process models) was quite weak: yields were available from Ullmann's Encyclopedia of Industrial Chemistry (2006) for only two of the three processes; the third was estimated with

Table 1 Composition of the tobacco flavor as described in Naegeli (1978)

Chemical	Content (mg/g flavor)	Actual data available in ecoinvent
Methylcyclopentenolone	2	No
Ethyl acetate	2	Yes
Ethyl anisate	4	No
Butyl formate	4	No
Theaspiran	5	No
Cinnamaldehyde	7	No
Capric aldehyde	10	No
Vanillin	10	No
Amyl salicylate	10	No
C14-aldehyde	10	No
Ethylvanillin	20	No
Heliotropin	20	No
Propyl acetate	25	Yes
Amyl formate	25	No
Isoamyl acetate	25	Yes
Coumarin	60	No
Ethyl butyrate	75	No
Benzaldehyde	110	Yes
Benzyl benzoate	250	No
Ethanol	326	Yes

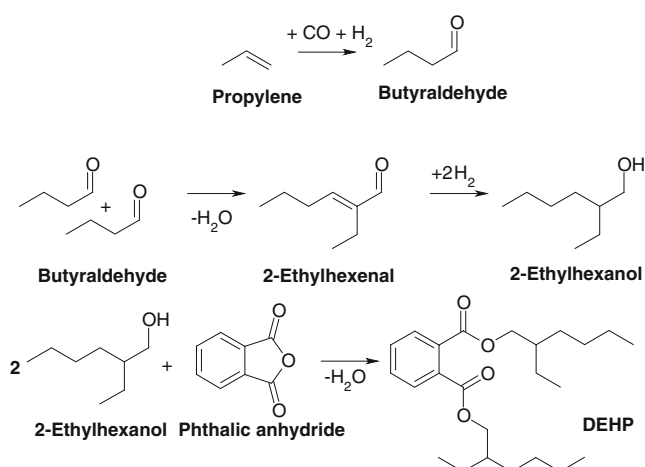
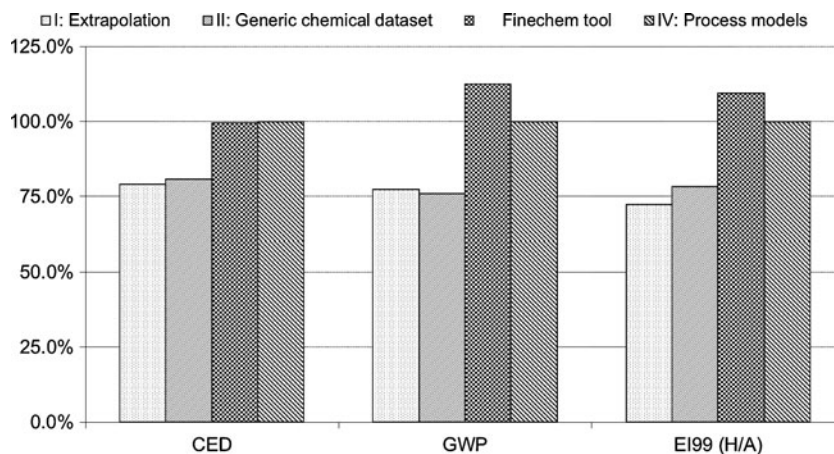


Fig. 2 Synthesis route of DEHP

Fig. 3 Relative impact assessment results of soft PVC using the different inventory modeling alternatives (soft PVC with process model estimation of DEHP=100 %). Tiered approach (V) identical to process models (IV) in this case



the standard estimations described in the [Electronic supplementary material](#). The reactions required no solvents, which reduced data requirements. No data were available for emissions to air and water or for electricity and steam use. The standard estimations were therefore used for energy, waste treatments, and emissions. As both the aldol condensation and the esterification require sustained reaction temperatures of 150–200°C, the standard estimations for energy use described above were likely too low. However, determining energy needs requires detailed information not just on the reaction process but also on equipment and plant conditions (Sziijarto et al. 2008). There is only a weak relationship between thermodynamic and actual energy requirements, as losses and equipment operation often make up the majority of energy expenditures. Thermodynamic requirements can therefore be a drastic underestimation of actual demands (Wernet et al. 2010). This was thus an example where lack of data led to a likely underestimation of the actual impacts, as there was a high probability the reactions had higher energy requirements than the average chemical reaction considered for the standard estimations.

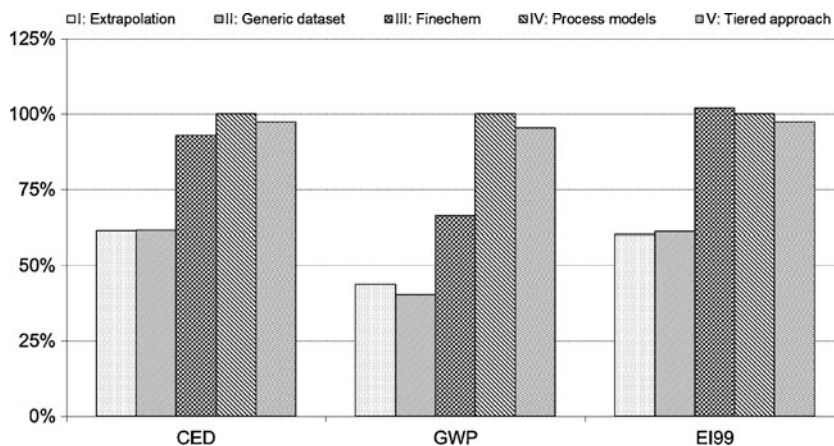
The analysis of the soft PVC scenario revealed that the first two approaches, namely extrapolation of DEHP data from PVC production and approximation with the generic

datasets “organic chemicals” from ecoinvent, lead to similar results, with regard to all impact LCIA methods considered. These results are significantly lower than the process model approach (approach 4) and the Finechem tool (approach 3). The scores of the latter two approaches are again comparable. As this was a mixture of only two chemicals, the tiered approach resulted in a recommendation of process models for both chemicals, so the results of approach V are in this case identical to the results of approach IV. The overall time requirement for the process model approach was in the order of a few days for data acquisition and inventory generation, while the time requirement for the other approaches was less than an hour.

3.2 Tobacco flavor

Again, the application of approaches 1–3 was simple and fast. For the sake of this study, the process model approach was applied to all 15 substances for which no ecoinvent data were available to allow a comparison of all approaches. The results can be seen in Fig. 4. They show that using the generic dataset for organic chemicals (approach 2) again leads to similar results as the extrapolation approach (approach 1), where the results for the known chemicals were

Fig. 4 Relative impact assessment results of the tobacco flavor using the different inventory modeling alternatives (process model approach=100 %)



scaled up to represent the whole mixture. The results of the Finechem tool (approach 3), the process models (approach 4), and the tiered approach (approach 5) for the full mixture of 20 chemicals are significantly higher. For the CED and EI99, the Finechem tool and the process models come to very similar conclusions, with only a few percent differences. For the GWP, the Finechem tool leads to only 70 % of the impacts compared to the process models. In this case, the tiered approach correctly helped to identify relevant chemicals and to selectively apply the process models. Process models for only five substances were necessary to apply the tiered approach in the case of the GWP: benzyl benzoate, ethyl butyrate, coumarin, heliotropin, and ethyl-vanillin. Four of these five were required each for the estimations of the CED and of the EI99 scores. The results of the tiered approach estimation differ from the process models by only a few percentage points. Applying the process models for all 15 chemicals required a time effort of about 3 weeks, so applying only the tiered approach would have saved roughly 2 weeks of work with minimal difference in results.

4 Discussion

In both case studies, the first two approaches of data extrapolation and approximation with a generic dataset lead to comparable results. For the extrapolation, the explanation is that in both case studies, the chemicals for which data were available were simpler to produce than the others, so the impacts were lower. As for the approximation, the generic dataset is in fact only a mix of a set of some of the most basic organic chemicals for which data are present in ecoinvent. Using this dataset to fill gaps therefore affects all results in the same way—they are shifted towards the average, generic result for these very simple organic chemicals. In most cases, using the generic dataset approximation will therefore underestimate the real impacts as well.

Using such proxy datasets is surely a better procedure than neglecting the missing inventories completely, but the quality of the results depends on the similarity of the chemicals with known and unknown LCI data. For example, in the plasticizer case study, it makes little sense from a synthesis perspective to substitute DEHP with PVC, since the production routes and emission profiles differ drastically. Similarly, the problem of the generic dataset is that it is a mix of very common and basic chemicals. If the chemical that is approximated is more complex in production, as in the cases investigated here, the generic dataset will underestimate the impacts of the production. The introduction of a scaling factor might be helpful in some scenarios, but any scaling factor will be an approximation of the complexity of the production. Given that these approaches are usually

employed only when data on the production are scarce, any scaling factor will be arbitrary. The generic dataset for “organic chemicals” is therefore only of limited usefulness as a proxy for chemicals. In specific cases, where high-quality data are available on a very similar production, substitution with that specific dataset may be an adequate option, but it takes knowledge of the production routes and in-depth knowledge of chemical synthesis to identify reasonable substitution candidates. In these cases, the existing data may be modified to describe the production of the desired chemical, but such scenarios are rare (e.g., this was the case for butyraldehyde in this study). It should be noted that adequate substitution options cannot be identified by similarity in name, structure, or application, as these are not necessarily related to similarities in the production.

MSMs can offer a fast and simple solution for impact estimations, and they have been shown to estimate chemical impacts with acceptable error margins for screening purposes (Wernet et al. 2009). They are however constrained in their application range by the training data used to create them. Furthermore, their “black box” approach means that introspection into the process chain and an identification of specific impact sources along the production chain are not possible. This may be an issue in cases where more detailed knowledge about the processes is required. Process-based models have the potential to accurately reflect the environmental burdens of a production and to identify specific impact sources, but they are susceptible to large errors if too few data are available to form the basis of the estimation. As seen in the plasticizer scenario, data on emissions and energy use are often difficult to acquire. The standard estimations of energy use and emissions recommended by ecoinvent and used in the process models for DEHP in this article refer to basic chemical production. Advanced and fine chemical productions generally require more specialized equipment and more complex reactions, leading to higher energy uses and more diverse emissions. The use of more complex solvents is also common. The reliance on standard estimations developed on basic chemical productions may thus cause process models to underestimate impacts of chemical productions, especially in impact categories related to emissions or energy use.

Each of these methods delivers results with an inherent uncertainty. Therefore, all methods are applicable only for screening LCAs and should be combined with a sensitivity analysis to determine the contribution of estimated impacts to the overall environmental burden. Inventories with a high relevance should not be estimated but based on actual data gathering at a production site or data published in literature.

The time effort required for the different models is noteworthy as well. While extrapolation and use of the generic dataset require little preparation time, the results of these methods are fundamentally questionable in cases where

inventories of complex chemicals are established. The results of high-quality process models are likely to be the most reliable of all, but they require access to detailed production data over the whole production chain, which may include one or more precursor productions, and a significant time investment. If the full production data are not available, the results can suffer from large uncertainties and may not be considered worth the effort compared to the first two methods. The recent introduction of MSMs can fill this gap by providing more reliable estimations in scenarios where data are especially scarce. These estimated data may then be used for screening assessments, and a sensitivity analysis may be employed to analyze the potential importance of the chemical production to the overall life cycle under discussion. MSMs like the Finechem tool are therefore valuable indicators of whether a production is important enough to warrant the full-time effort for a process model approach, including data gathering. This is demonstrated in the tiered approach: in the flavor scenario, the tiered approach led to four (CED and EI99) or five (GWP) chemicals being assessed with process models. That means that the effort required for an analysis of the 15 chemicals on a process level could be reduced by two thirds with little effect on the result quality. This combination of Finechem and process models thus offers the potential for data quality of the process models with a significant reduction of the efforts required for the overall estimation of chemical mixtures. Table 2 gives an overview of the different methods and their capabilities.

5 Conclusions

While extrapolation or approximation of chemicals, especially in large numbers as is the case in mixtures, with similar or generic datasets may seem tempting due to the low requirements on time or data, these approaches cannot generally be recommended given the results of the two case studies presented here. Depending on the amount of production data available, either a process model approach or the use of the Finechem tool is preferable in most cases. These two approaches complement each other well, as they have different data requirements. Process models suffer from the high data requirements as shown in Table 2, and

they may underestimate the complexities of fine chemical production if standard estimations for basic chemicals are used. Unfortunately, generic models with average values for energy use and emissions in fine chemical production do not yet exist. On the other hand, the Finechem tool is especially useful in cases where few data are available. The Finechem tool is therefore a useful addition to the toolbox of LCA practitioners and offers a novel solution to a problem commonly encountered in LCAs. If the Finechem tool or other approximation methods indicate a significant relevance of the estimated inventories, additional data gathering projects should be initiated to allow a thorough assessment with reduced uncertainties. This is exemplified in the tiered approach shown in Fig. 1. The results of this tiered approach show that a combination of Finechem tool and process models can reduce the workload required while allowing users to improve the results with targeted data acquisition for process models on the most relevant components of a chemical mixture. In case study 2, 2 weeks of estimation work could be avoided by the use of the tiered approach.

6 Recommendation and perspectives

Extrapolation and using generic datasets do not differentiate between different chemicals and can generally not be recommended for the estimation of production impacts of complex chemicals. Depending on the overall relevance of the chemical mixture in an LCA study, different approaches can be recommended. If the overall relevance is rather low, the Finechem model offers a fast and simple solution. If the relevance is high, process models are recommended, as they offer the possibility of including detailed production data that should be gathered in such cases. The tiered approach allows LCA practitioners to benefit from the advantages of both approaches and should be considered especially if a large number of chemicals need to be assessed.

Still, both approaches, the Finechem tool and process models, need to be further improved. The neural network models used in Finechem are generic approximators, and their usefulness to model other industry sectors is an intriguing research topic. In the field of chemicals, the tool is currently limited to organic chemicals, and some chemical classes are not covered by the tool. The capabilities of the

Table 2 Overview of the different estimation methods

Method	1: Extrapolation	2: Generic dataset	3: Finechem	4: Process models
Data requirements	Mass, inventory data of similar chemical	Mass, inventory data of generic dataset	Mass, molecular structure	Mass, synthesis route, detailed process data
Time requirements per chemical production step	Minutes	Minutes	<1 h	Hours–days
Chemical knowledge required	Low	Low	Medium	High

neural networks can be extended once training data from chemical classes that are currently not covered become available. Concerning the process model approach, standard estimations more suited to fine chemical production and advanced chemistry would allow a better estimation of more complex chemicals. In chemistry, the reaction type often defines the production parameters more than the specific product and the development of reaction-specific models and estimations would be a significant step towards increasing the quality of process models in chemical production while reducing the requirements and efforts for their application. Such process-specific models would also be able to interact with chemical process simulation and flow-sheeting software and other models beyond the scope of LCA as well.

Acknowledgments We thank Jürgen Sutter for his help in creating the process model inventories of the chemicals in the flavor scenario.

References

- Bretz R, Frankhauser P (1997) Life-cycle assessment of chemical production processes: a tool for ecological optimization. *Chimia* 51:213–217
- Capello C, Hellweg S, Badertscher B, Betschart H, Hungerbühler K (2007) Environmental assessment of waste-solvent treatment options—part 1: the ecosolvent tool—environmental assessment of waste-solvent treatment options. *J Ind Ecol* 11:26–38
- Curzons AD, Jiménez-Gonzalez C, Duncan AL, Constable DJC, Cunningham VL (2007) Fast life cycle assessment of synthetic chemistry (FLASC) tool. *Int J Life Cycle Assess* 12:272–280
- EU JRC (2008) Bis(2-ethylhexyl) phthalate (DEHP). Summary risk assessment report. European Chemicals Bureau, Ispra
- Geisler G, Hofstetter TB, Hungerbühler K (2004) Production of fine and speciality chemicals: procedure for the estimation of LCIs. *Int J Life Cycle Assess* 9:101–113
- Goedkoop M, Spriensma R (2000) The Eco-indicator 99: a damage orientated method for life cycle impact assessment. PRé Consultants, Amersfoort
- Hischier R, Hellweg S, Capello C, Primas A (2004) Establishing life cycle inventories of chemicals based on differing data availability. *Int J Life Cycle Assess* 10:59–67
- IPCC (2007) Fourth report. Climate change 2007
- Jiménez-Gonzalez C, Kim S, Overcash M (2000) Methodology for developing gate-to-gate life cycle inventory information. *Int J Life Cycle Assess* 5:153–159
- Jiménez-Gonzalez C, Curzons AD, Constable DJC, Cunningham VL (2004) Cradle-to-gate life cycle inventory and assessment of pharmaceutical compounds. *Int J Life Cycle Assess* 9:114–121
- Köhler A, Hellweg S, Recan E, Hungerbühler K (2007) Input-dependent life-cycle inventory model of industrial wastewater-treatment processes in the chemical sector. *Environ Sci Technol* 41:5515–5522
- Naegeli P (1978) In: United States Patent Office (ed) 4-(2,6,6-Trimethyl-2-cyclohexen-1-ylidene)-butan-2-ol. Givaudan Corp., Clifton. Patent number 4072719
- NREL (2008) U.S. Life-cycle inventory database. National renewable energy laboratory
- PlasticsEurope (2005) Eco-profiles of the European plastics industry
- Saling P, Kicherer A, Dittrich-Krämer B, Wittlinger R, Zombik W, Schmidt I, Schrott W, Schmidt S (2002) Eco-efficiency analysis by BASF: the method. *Int J Life Cycle Assess* 7:203–218
- Swingler K (1996) Applying neural networks. Academic, London
- Szijjarto A, Papadokonstadakis S, Fischer U, Hungerbühler K (2008) Bottom-up modeling of the steam consumption in multipurpose chemical batch plants focusing on identification of the optimization potential. *Ind Eng Chem Res* 47:7323–7334
- The ecoinvent centre (2008a) Ecoinvent 2.01 database
- The ecoinvent centre (2008b) Ecoinvent 2.01 reports, chapter 1: overview and methodology
- Thieme Chemistry (1995) Römpp online <http://www.roempp.com>. Accessed 2010
- Ullmann's Encyclopedia of Industrial Chemistry (2006) 7th electronic edition. <http://www.mrw.interscience.wiley.com/ueic/>, Wiley-VCH, Weinheim. Accessed 2010
- Van der Vorst G, Dewulf J, Aelterman W, De Witte B, Van Langenhove H (2009) Assessment of the integral resource consumption of individual chemical production processes in a multipurpose pharmaceutical production plant: a complex task. *Ind Eng Chem Res* 48:5344–5350
- VDI (1997) VDI-richtlinie 4600: cumulative energy demand, terms, definitions, methods of calculation. Verein Deutscher Ingenieure, Düsseldorf
- Wernet G, Hellweg S, Fischer U, Papadokonstadakis S, Hungerbühler K (2008) Molecular-structure-based models of chemical inventories using neural networks. *Environ Sci Technol* 42:6717–6722
- Wernet G, Papadokonstadakis S, Hellweg S, Hungerbühler K (2009) Bridging data gaps in environmental assessments: modeling impacts of fine and basic chemical production. *Green Chem* 11(11):1826–1831
- Wernet G, Conradt S, Isenring HP, Jiménez-Gonzalez C, Hungerbühler K (2010) Life cycle assessment of fine chemical production: a case study of pharmaceutical synthesis. *Int J Life Cycle Assess* 15:294–303